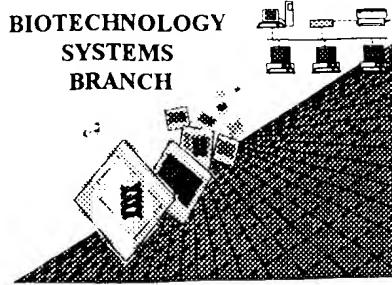


590
12/14/97

RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/893,666

Source: O1PE

Date Processed by STIC: 12/6/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§ 1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>09/893,666</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input type="checkbox"/> Wrapped Nucleic Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped	
	Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input checked="" type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001
TIME: 10:36:14

Input Set : A:\210217US0.ST25.txt
Output Set: N:\CRF3\12062001\I893666.raw

3 <110> APPLICANT: YAMASHITA, ICHIRO
 5 <120> TITLE OF INVENTION: High estrogen-sensitive medaka fish
 7 <130> FILE REFERENCE: 210217US-620-7249-0
 9 <140> CURRENT APPLICATION NUMBER: US 09/893,666
 10 <141> CURRENT FILING DATE: 2001-06-29
 12 <150> PRIOR APPLICATION NUMBER: JP 2000-247729
 13 <151> PRIOR FILING DATE: 2000-08-17
 15 <160> NUMBER OF SEQ ID NOS: 7
 17 <170> SOFTWARE: PatentIn version 3.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 2764
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Oryzias latipes
 24 <220> FEATURE:
 25 <221> NAME/KEY: CDS
 26 <222> LOCATION: (211)..(1935)
 27 <223> OTHER INFORMATION:
 30 <400> SEQUENCE: 1
 31 qtctcgctgc tagatgcctg tcagggcaggc agagaggaag cagcccggtgt tgccgcac 60
 33 atctgaggat gattcatgag taagagacag agctcggtgc agatcaggca gctgttcgga 120
 35 ccagcaactca gatccaggat cagcccagcc tccctcagagc tggagaccct ctccccacct 180
 37 cgcctctcgc cccgtgaccc cctcggtgac atg tac cct gaa gag agc cgg ggt 234
 38 Met Tyr Pro Glu Glu Ser Arg Gly
 39 1 5
 41 tct gga ggg gtg gct gct gtg gac ttt ttg gaa ggg acg tac gac tat 282
 42 Ser Gly Gly Val Ala Ala Val Asp Phe Leu Glu Gly Thr Tyr Asp Tyr
 43 10 15 20
 45 qcc gcc ccc aac oct gcc acg act ccc ctt tac acg cag tcc agc acc 330
 46 Ala Ala Pro Asn Pro Ala Thr Thr Pro Leu Tyr Ser Gln Ser Ser Thr
 47 25 30 35 40
 49 ggc tac tac tct gct ccc ctg gaa aca aac gga ccc ccc tca gaa ggc 378
 50 Gly Tyr Ser Ala Pro Leu Glu Thr Asn Gly Pro Pro Ser Glu Gly
 51 45 50 55
 53 agt ctg cag tcc ctg ggc agt ggg ccg acg acg cct ctg gtg ttt gtg 426
 54 Ser Leu Gln Ser Leu Gly Ser Gly Pro Thr Ser Pro Leu Val Phe Val
 55 60 65 70
 57 ccc tcc agc ccc aga ctc agt ccc ttt atg cat cca ccc agc cac cac 474
 58 Pro Ser Ser Pro Arg Leu Ser Pro Phe Met His Pro Pro Ser His His
 59 75 80 85
 61 tat ctg gaa acc act tcc acg ccc gtt tac aga tcc agc cac cag gga 522
 62 Tyr Leu Glu Thr Ser Thr Pro Val Tyr Arg Ser Ser His Gln Gly
 63 90 95 100
 65 gcc tcc agg gag gac cag tgc ggc tcc cgg gag gac acg tgc agc ctg 570
 66 Ala Ser Arg Glu Asp Gln Cys Gly Ser Arg Glu Asp Thr Cys Ser Leu
 67 105 110 115 120
 69 ggg gag tta ggc gcc gga gcc ggg gct ggg ggg ttt gag atg gcc aaa 618
 70 Gly Glu Ieu Gly Ala Gly Ala Gly Phe Glu Met Ala Lys

P6

Does Not Comply
Correct Amended Sequence

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001
TIME: 10:36:14

Input Set : A:\210217US0.ST25.txt
Output Set: N:\CRF3\12062001\I893666.raw

71	125	130	135	
73	gac acg cgt ttc tgc gcc gtg tgc agc gac tac gcc tct ggg tac cac			666
74	Asp Thr Arg Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His			
75	140	145	150	
77	tat ggg gtg tgg tct tgt gag ggc tgc aag gcc ttc ttc aag agg agc			714
78	Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser			
79	155	160	165	
81	atc cag ggt cac aat gac tat atg tgc cca gcg acc aat cag tgc act			762
82	Ile Gln Gly His Asn Asp Tyr Met Cys Pro Ala Thr Asn Gln Cys Thr			
83	170	175	180	
85	att gac aga aat cggtt aag aac tgc cag gct tgt cgt ctt agg aag			810
86	Ile Asp Arg Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys			
87	185	190	195	200
89	tgt tac gaa gtg gga atg atg aaa ggc ggt gtg cgc aag gac cgc att			858
90	Cys Tyr Glu Val Gly Met Met Lys Gly Gly Val Arg Lys Asp Arg Ile			
91	205	210	215	
93	cgc att tta cggtt cgt gac aaa cggtt cgg aca ggc gtt ggt gat gga gac			906
94	Arg Ile Leu Arg Arg Asp Lys Arg Arg Thr Gly Val Gly Asp Gly Asp			
95	220	225	230	
97	aag gtt gta aag ggt cag gag cat aaa acg gtg cat tat gat gga agg			954
98	Lys Val Val Lys Gly Gln Glu His Lys Thr Val His Tyr Asp Gly Arg			
99	235	240	245	
101	aaa cgc agc agc aca gga gga gga gga gga gga gga aga ctg			1002
102	Lys Arg Ser Ser Thr Gly Gly Gly Gly Gly Gly Arg Leu			
103	250	255	260	
105	tct gtg acc agc ata cct cct gag cag gtg ctg ctc ctc ctt cag ggc			1050
106	Ser Val Thr Ser Ile Pro Pro Glu Gln Val Leu Leu Leu Gln Gly			
107	265	270	275	280
109	gcc gag ccc ccg ata ctc tgc tgc cgt cag aag ttg agc cga ccg tac			1098
110	Ala Glu Pro Pro Ile Leu Cys Ser Arg Gln Lys Leu Ser Arg Pro Tyr			
111	285	290	295	
113	acc gag gtc acc atg atg acc ctg ctc acc agc atg gca gac aag gag			1146
114	Thr Glu Val Thr Met Met Thr Leu Leu Thr Ser Met Ala Asp Lys Glu			
115	300	305	310	
117	ctg gtc cac atg atc gcc tgg gcc aag aac ctc cca ggt ttt ctg cag			1194
118	Leu Val His Met Ile Ala Trp Ala Lys Lys Leu Pro Gly Phe Leu Gln			
119	315	320	325	
121	ctg tcc ctg cac gat cag gtg ctg ctg gag agc tcg tgg ctg gag			1242
122	Leu Ser Leu His Asp Gln Val Leu Leu Leu Glu Ser Ser Trp Leu Glu			
123	330	335	340	
125	gtg ctc atg atc ggc ctc att tgg agg tcc atc cac tgt ccc ggg aag			1290
126	Val Leu Met Ile Gly Leu Ile Trp Arg Ser Ile His Cys Pro Gly Lys			
127	345	350	355	360
129	ctc atc ttt gca caa gac ctc atc ctg gac agg aat gag gga gac tgc			1338
130	Leu Ile Phe Ala Gln Asp Leu Ile Leu Asp Arg Asn Glu Gly Asp Cys			
131	365	370	375	
133	gtg gaa ggc atg acg gag atc ttc gac atg ctg ctg gcc act gct tcc			1386
134	Val Glu Gly Met Thr Glu Ile Phe Asp Met Leu Leu Ala Thr Ala Ser			
135	380	385	390	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001

TIME: 10:36:14

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\12062001\I893666.raw

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001
TIME: 10:36:14

Input Set : A:\210217US0.ST25.txt
Output Set: N:\CRF3\12062001\I893666.raw

```

211 qcaattcctg tttctatTTT 2764
214 <210> SEQ ID NO: 2
215 <211> LENGTH: 575
216 <212> TYPE: PRT
217 <213> ORGANISM: Oryzias latipes
219 <400> SEQUENCE: 2
221 Met Tyr Pro Glu Glu Ser Arg Gly Ser Gly Val Ala Ala Val Asp
222 1 5 10 15
225 Phe Leu Glu Gly Thr Tyr Asp Tyr Ala Ala Pro Asn Pro Ala Thr Thr
226 20 25 30
229 Pro Leu Tyr Ser Gln Ser Ser Thr Gly Tyr Tyr Ser Ala Pro Leu Glu
230 35 40 45
233 Thr Asn Gly Pro Pro Ser Glu Gly Ser Leu Gln Ser Leu Gly Ser Gly
234 50 55 60
237 Pro Thr Ser Pro Leu Val Phe Val Pro Ser Ser Pro Arg Leu Ser Pro
238 65 70 75 80
241 Phe Met His Pro Pro Ser His His Tyr Leu Glu Thr Thr Ser Thr Pro
242 85 90 95
245 Val Tyr Arg Ser Ser His Gln Gly Ala Ser Arg Glu Asp Gln Cys Gly
246 100 105 110
249 Ser Arg Glu Asp Thr Cys Ser Leu Gly Glu Leu Gly Ala Gly Ala Gly
250 115 120 125
253 Ala Gly Gly Phe Glu Met Ala Lys Asp Thr Arg Phe Cys Ala Val Cys
254 130 135 140
257 Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly
258 145 150 155 160
261 Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met
262 165 170 175
265 Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Arg Asn Arg Arg Lys Ser
266 180 185 190
269 Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys
270 195 200 205
273 Gly Gly Val Arg Lys Asp Arg Ile Arg Ile Leu Arg Arg Asp Lys Arg
274 210 215 220
277 Arg Thr Gly Val Gly Asp Gly Asp Lys Val Val Lys Gly Gln Glu His
278 225 230 235 240
281 Lys Thr Val His Tyr Asp Gly Arg Lys Arg Ser Ser Thr Gly Gly Gly
282 245 250 255
285 Gly Gly Gly Gly Gly Arg Leu Ser Val Thr Ser Ile Pro Pro Glu
286 260 265 270
289 Gln Val Ile Leu Leu Leu Gln Gly Ala Glu Pro Pro Ile Leu Cys Ser
290 275 280 285
293 Arg Gln Lys Leu Ser Arg Pro Tyr Thr Glu Val Thr Met Met Thr Leu
294 290 295 300
297 Leu Thr Ser Met Ala Asp Lys Glu Leu Val His Met Ile Ala Trp Ala
298 305 310 315 320
301 Lys Lys Ile Pro Gly Phe Leu Gln Leu Ser Leu His Asp Gln Val Leu
302 325 330 335
305 Leu Leu Glu Ser Ser Trp Leu Glu Val Leu Met Ile Gly Leu Ile Trp

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001
 TIME: 10:36:14

Input Set : A:\210217US0.ST25.txt
 Output Set: N:\CRF3\12062001\I893666.raw

306	340	345	350	
309 Arg Ser Ile His Cys Pro Gly Lys Leu Ile Phe Ala Gln Asp Leu Ile				
310	355	360	365	
313 Leu Asp Arg Asn Glu Gly Asp Cys Val Glu Gly Met Thr Glu Ile Phe				
314	370	375	380	
317 Asp Met Leu Leu Ala Thr Ala Ser Arg Phe Arg Val Leu Lys Leu Lys				
318	385	390	395	400
321 Pro Glu Glu Phe Val Cys Leu Lys Ala Ile Ile Leu Leu Asn Ser Gly				
322	405	410	415	
325 Ala Phe Ser Phe Cys Thr Gly Thr Met Glu Pro Leu His Asn Ser Ala				
326	420	425	430	
329 Ala Val Gln Ser Met Leu Asp Thr Ile Thr Asp Ala Leu Ile His Tyr				
330	435	440	445	
333 Ile Ser Gln Ser Gly Tyr Leu Ala Gln Glu Gln Ala Arg Arg Gln Ala				
334	450	455	460	
337 Gln Leu Leu Leu Leu Ser His Ile Arg His Met Ser Asn Lys Gly				
338	465	470	475	480
341 Met Glu His Leu Tyr Ser Met Lys Cys Lys Asn Lys Val Pro Leu Tyr				
342	485	490	495	
345 Asp Leu Leu Leu Glu Met Leu Asp Ala His Arg Leu His His Pro Val				
346	500	505	510	
349 Arg Ala Pro Gln Ser Leu Ser Gln Val Asp Arg Asp Pro Pro Ser Thr				
350	515	520	525	
353 Ser Ser Gly Gly Gly Ile Ala Pro Gly Ser Ile Ser Ala Ser Arg				
354	530	535	540	
357 Gly Arg Ile Glu Ser Pro Ser Arg Gly Pro Phe Ala Pro Ser Val Leu				
358	545	550	555	560
361 Gln Tyr Gly Gly Ser Arg Pro Asp Cys Thr Pro Ala Leu Gln Asp				
362	565	570	575	
365 <210> SEQ ID NO: 3				
366 <211> LENGTH: 20				
367 <212> TYPE: DNA				
368 <213> ORGANISM: ARTIFICIAL SEQUENCE				
370 <220> FEATURE:				
371 <223> OTHER INFORMATION: SYNTHETIC DNA				
373 <400> SEQUENCE: 3				
374 tcgggtgacat gtaccctgaa			20	
377 <210> SEQ ID NO: 4				
378 <211> LENGTH: 20				
379 <212> TYPE: DNA				
380 <213> ORGANISM: ARTIFICIAL SEQUENCE				
382 <220> FEATURE:				
383 <223> OTHER INFORMATION: SYNTHETIC DNA				
385 <400> SEQUENCE: 4				
386 ctgttgtqctc agtcttqaag			20	
389 <210> SEQ ID NO: 5				
390 <211> LENGTH: 20				
391 <212> TYPE: DNA				
392 <213> ORGANISM: ARTIFICIAL SEQUENCE				

07/893, 666 6

<210> 6

<211> 20

<212> DNA

<213> ARTIFICIAL DNA

<400> 6

gtaggaggc ataaagaggg

see item 10 on Error Summary Sheet

20

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/893,666

DATE: 12/06/2001

TIME: 10:36:15

Input Set : A:\210217US0.ST25.txt

Output Set: N:\CRF3\12062001\I893666.raw

L:404 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6

L:406 M:258 W: Mandatory Feature missing, <220> FEATURE:

L:406 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION: